

SRT Evaluation of AIRS Version-6.02 and Version-6.02 AIRS Only (6.02 AO) Products

Joel Susskind, Lena Iredell, Gyula Molnar, and John Blaisdell

1. Introduction

Version-6 contains a number of significant improvements over Version-5. This report compares Version-6 products resulting from the advances listed below to those from Version-5.

1. Improved methodology to determine skin temperature (T_s) and spectral emissivity (ϵ_v).
2. Use of Neural-net start-up state (Bill Blackwell).
3. Improvements which decrease the spurious negative Version-5 trend in tropospheric temperatures (Eric Maddy).
4. Improved QC methodology. Version-6 uses separate QC thresholds optimized for Data Assimilation (QC=0) and Climate applications (QC=0,1) respectively.
5. Channel-by-channel clear-column radiances $\bar{R}_t$ QC flags.
6. Improved cloud parameter retrieval algorithm (Evan Manning, Van Dang, John Blaisdell).
7. Improved OLR RTA (Gyula Molnar).

Our evaluation compared V6.02 and V6.02 AIRS Only (V6.02 AO) Quality Controlled products with those of Version-5.0. In particular we evaluated surface skin temperature T_s , surface spectral emissivity ϵ_v , temperature profile $T(p)$, water vapor profile $q(p)$, OLR, OLR_{CLR} , effective cloud fraction $\alpha\epsilon$, and cloud cleared radiances $\bar{R}_t$. We conducted two types of evaluations. The first compared results on 7 focus days to collocated ECMWF truth. The seven focus days are: September 6, 2002; January 25, 2003; September 29, 2004; August 5, 2005; February 24, 2007; August 10, 2007; and May 30, 2010. In these evaluations, we show results for T_s , ϵ_v , $T(p)$, and $q(p)$ in terms of yields, and RMS differences and biases with regard to ECMWF. We also show yield trends as well as bias trends of these quantities relative to ECMWF truth. We also show yields and accuracy of channel by channel QC'd values of $\bar{R}_t$ for V6.02 and V6.02 AO. Version-5 did not contain channel by channel QC'd values of $\bar{R}_t$.

In the second type of evaluation, we compared V6.03 monthly mean Level-3 products to those of Version-5.0, for four different months: January, April, July, and October; in 3 different years 2003, 2007, and 2011. In particular, we compared V6.03 and V5.0 trends of $T(p)$, $q(p)$, $\alpha\epsilon$, OLR, and OLR_{CLR} computed based on results for these 12 time periods.

2. Validation Results Using Seven Focus Days

2.1 Surface Skin Temperature T_s and Surface Spectral Emissivity ϵ_v

The most significant difference in the retrieval methodology used in Version-6 and in Version-5 is the approach used to determine T_s and ϵ_v . Version-5 simultaneously retrieves T_s , longwave surface emissivity $\epsilon_{v\ell w}$, and shortwave surface emissivity ϵ_{vsw} using cloud cleared radiances $\bar{R}_\lambda$ for an ensemble of longwave and shortwave window channels. Following theoretical considerations, Version-6 simultaneously retrieves T_s and ϵ_{vsw} using $\bar{R}_\lambda$ for an ensemble of channels found only in the shortwave window regions. According to cloud clearing theory, the Version-6 approach allows for the determination of accurate values of T_s under

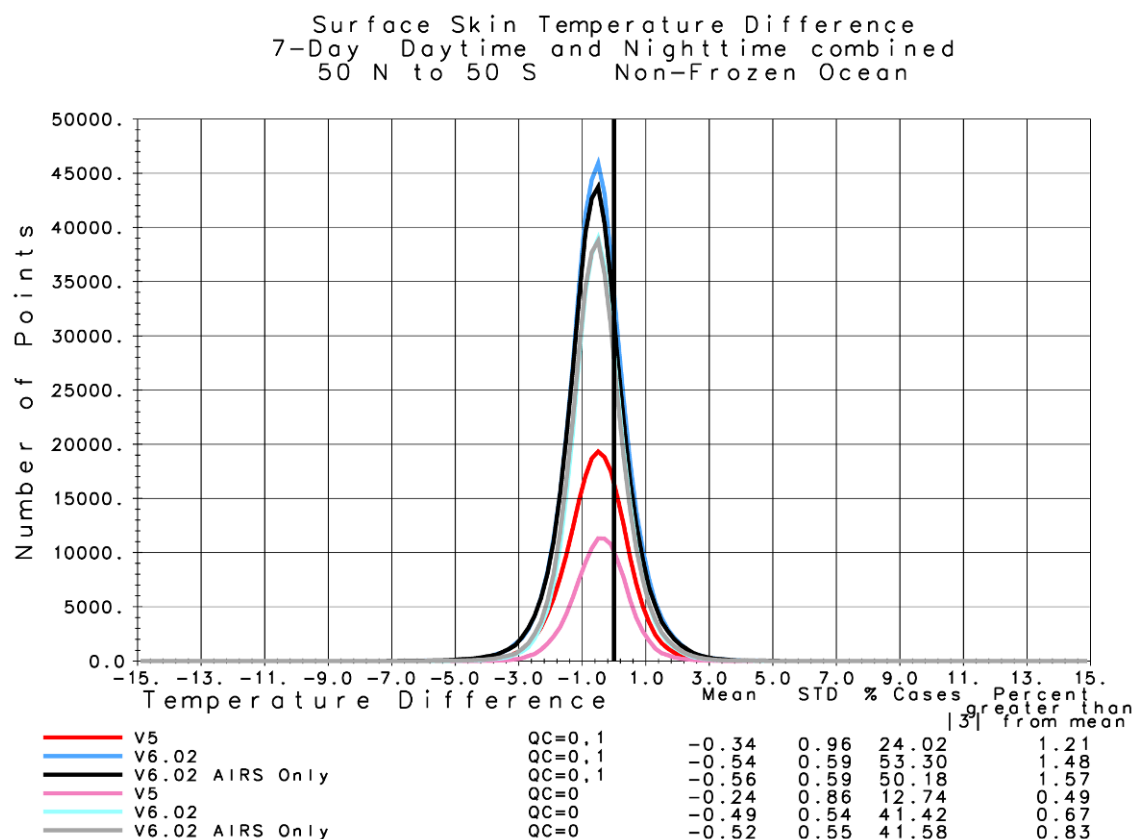


Figure 1

more stressing cloud conditions than does the Version-5 approach. Longwave surface spectral emissivity is solved for in Version-6 in a subsequent step, using $\bar{R}_\lambda$ in channels found only in the longwave window region, as well as values of T_s , $T(p)$, and $q(p)$ determined in previous steps. Figure 1 shows counts of all Quality Controlled Ocean Surface Skin Temperatures over the

latitude range, 50°N – 50°S, as a function of the difference between T_s and ECMWF “truth” for the 7-day evaluation period. Counts of Version-5 retrievals are shown in red and pink, Version-6.02 retrievals are shown in dark blue and light blue, and Version-6.02 AO retrievals are shown in black and gray. The lighter shade of each color shows counts of best quality T_s retrievals with QC=0 and the darker shade shows counts of both best and good quality T_s retrievals including cases with QC=0 or 1. Ocean T_s retrievals with QC=0 or 1 are the ensemble used to generate the Level-3 Oceanic SST product. Figure 1 also contains statistics for each set of retrievals showing the mean difference from ECMWF, the standard deviation of the ensemble differences, the percentage of all possible cases included in the Quality Controlled ensemble, and percentage of all accepted cases with absolute differences from ECMWF of more than 3K from the mean difference. Such cases are referred to as outliers.

Version-6.02 QC'd retrievals accept considerably more cases than Version-5 and have much lower standard deviations of the errors as well. In both ensembles, the percentage of outliers grows with loosening the QC thresholds as expected. Version-6.02 outliers with QC=0,1 are somewhat larger than Version-5, but the yield is more than twice as large. It is noteworthy that Version-6.02 retrievals with QC=0 have a much smaller percent outliers than does Version-5 retrievals with QC=0,1 along with a substantially higher yield. One point of slight concern in this figure is that the cold mean bias in Version-6.02 retrievals compared to ECMWF is somewhat larger than that of Version-5. Statistics of QC'd Version-6.02 AO retrievals are very similar to those of Version-6.02.

Figure 2 shows the spatial distribution of the seven daily differences of the Level-3 SST products from collocated ECMWF values for both Version-6.02 and Version-5. The values shown in a given grid box are the average values for that grid box of all accepted cases where the SST retrieval was accepted either at 1:30 AM or 1:30 PM. A grid point for which not a single value of QC'd SST was obtained for all 14 possible cases (seven days, twice daily) is shown in gray. Figure 2 represents the spatial coverage and accuracy of “pseudo seven day mean” Level-3 products. The seven days included in the figure are not consecutive, but the figure is very informative nonetheless.

The caption under each field indicates the mean difference of the Level-3 SST field from its own collocated ECMWF values, the spatial standard deviation over all grid points of the Level-3 differences, and the % of all possible grid points that have at least one accepted value over the seven day period (i.e., one not gray). The Version-6.02 Level-3 SST product has much better spatial coverage, with 98.35% of all oceanic grid points 60°N-60°S being filled, compared to Version-5 with only 92.24%. Moreover, there are large coherent spatial areas in which no Version-5 retrievals were accepted on any of the seven days. The spatial standard deviation of the Version-6.02 Level-3 SST product errors compared to ECMWF truth is also much smaller than that of Version-5, and the Version-6.02 area mean negative bias is also smaller than that of Version-5.

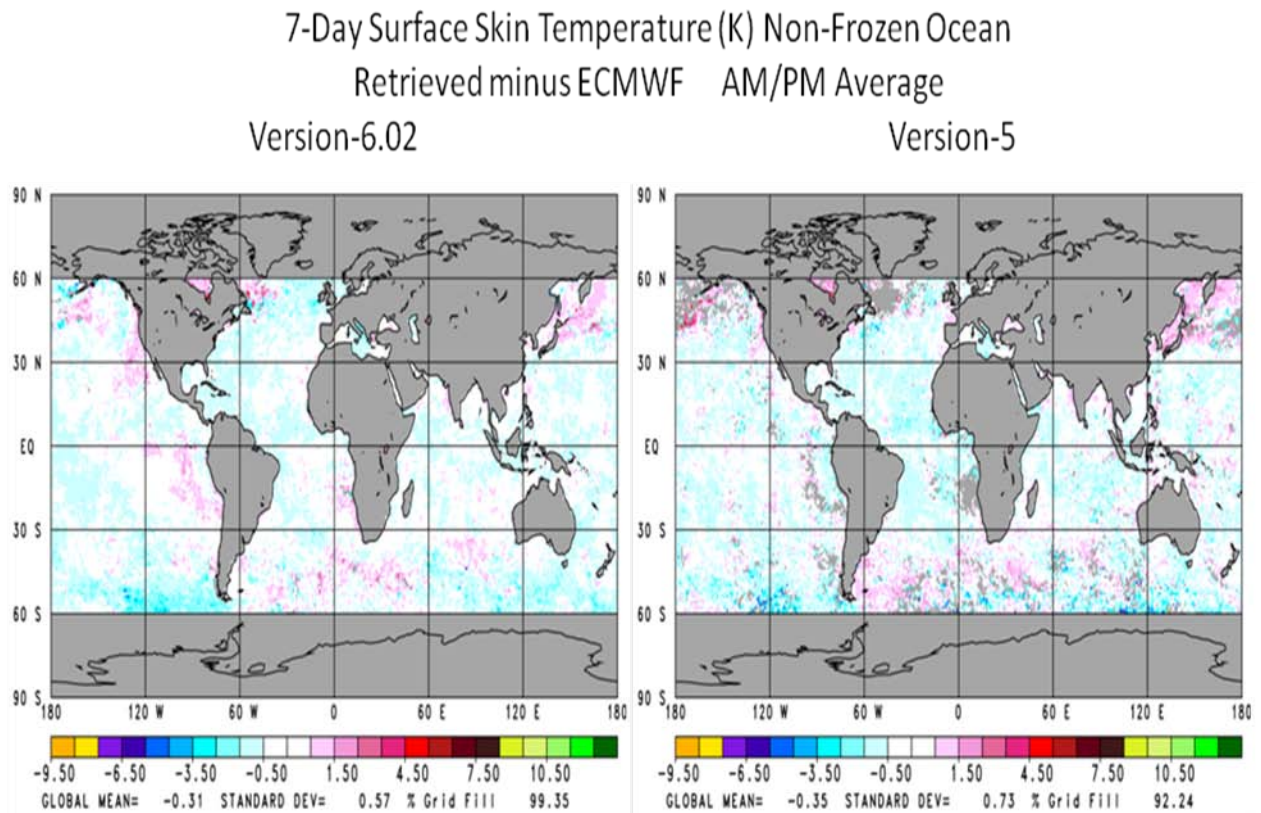


Figure 2

Figure 3 shows statistics related to retrieved ocean surface spectral emissivity ϵ_v as a function of satellite zenith angle for $\nu = 950 \text{ cm}^{-1}$ and $\nu = 2400 \text{ cm}^{-1}$. The channels are in the longwave and shortwave window regions respectively. In these figures, statistics are shown separately for AM orbits in dark colors, and PM orbits in light colors. Figures 3a and 3b show

the mean differences of retrieved values of ϵ_v from those values calculated using the Masuda ocean spectral emissivity model, which is taken as truth. In both the longwave and shortwave window regions, Version-6.02 (as well as Version-6.02 AO) retrieved ocean spectral emissivities as a function of satellite zenith angle are very close to the values expected using the Masuda emissivity model. Differences of Version-5 retrieved ϵ_v from ECMWF truth are much larger than those of Version-6 AO. Version-5 values of ϵ_v also show a large spurious feature during the day in the vicinity of a satellite zenith angle of -18.24 degrees, which is the viewing angle in which sunglint appears in the field of view. Figures 3c and 3d show the standard deviations of the retrieved values of ϵ_v from their mean values for the same two frequencies. These standard deviations are much smaller in Version-6.02 as compared to Version-5 indicating that the retrieved values are not only more accurate in Version-6.02 but considerably more stable as well. There is no appreciable difference between Version-6.02 and Version-6.02 AO in results related to ocean values of ϵ_v .

Surface spectral emissivity over land is not well known nor is it easily modeled. Nevertheless, land surface emissivity is not expected to change significantly from night to day. Therefore, it is useful to examine the characteristics of land surface emissivity determined at night minus those determined during the day. Figure 4 shows the 7-day mean value of the nighttime minus daytime retrieved Quality Controlled surface emissivity over land as a function of satellite zenith angle for 950 cm^{-1} and 2400 cm^{-1} . The (spurious) diurnal signal in Version-6.02 land surface emissivity is much smaller at both frequencies than found in Version-5. This is an indication that Version-6.02 land surface emissivities should be of higher quality than those of Version-6. Figure 5 shows the spatial distribution of the nighttime minus daytime seven-day mean land surface emissivities of 950 cm^{-1} and 2400 cm^{-1} for both Version-6.02 and Version-5. Some spurious day/night differences in land surface emissivity still exist in Version-6.02, particularly over the Sahara desert and Saudi Arabia at 2400 cm^{-1} , but these diurnal differences are much smaller than those found in Version-5.

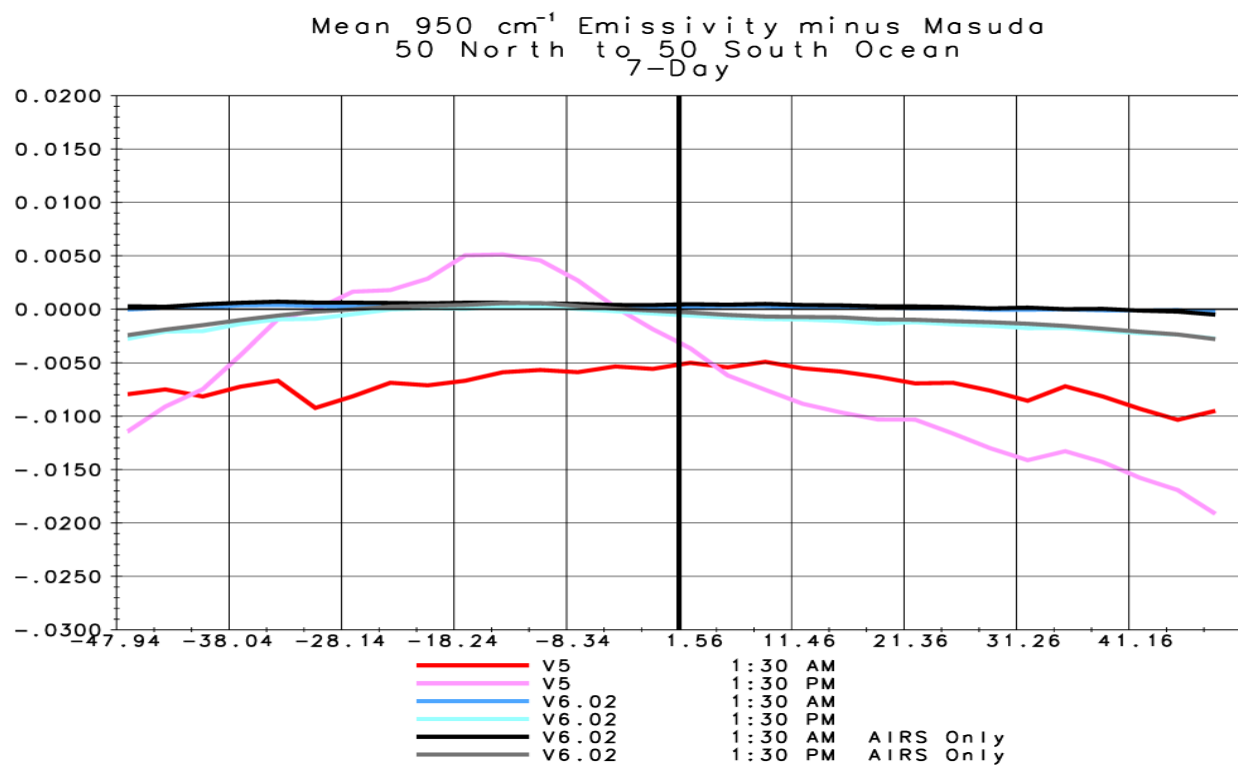


Figure 3a

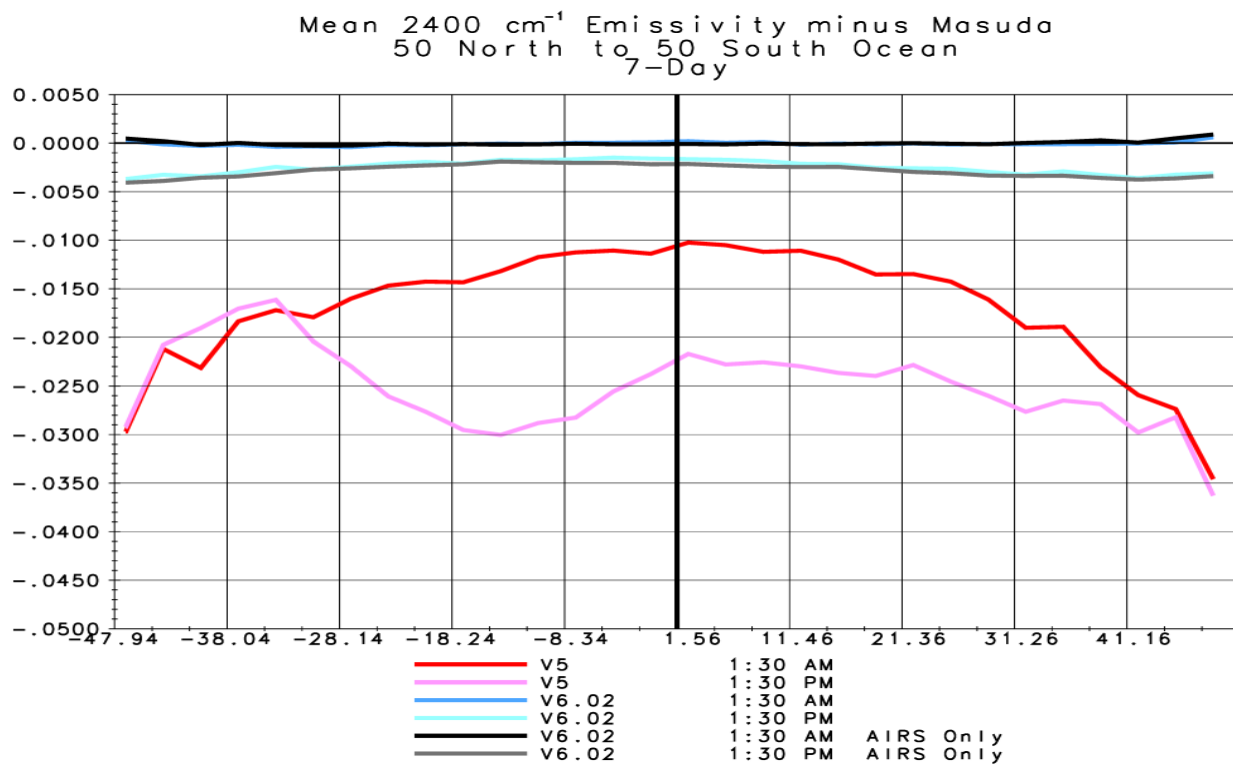


Figure 3b

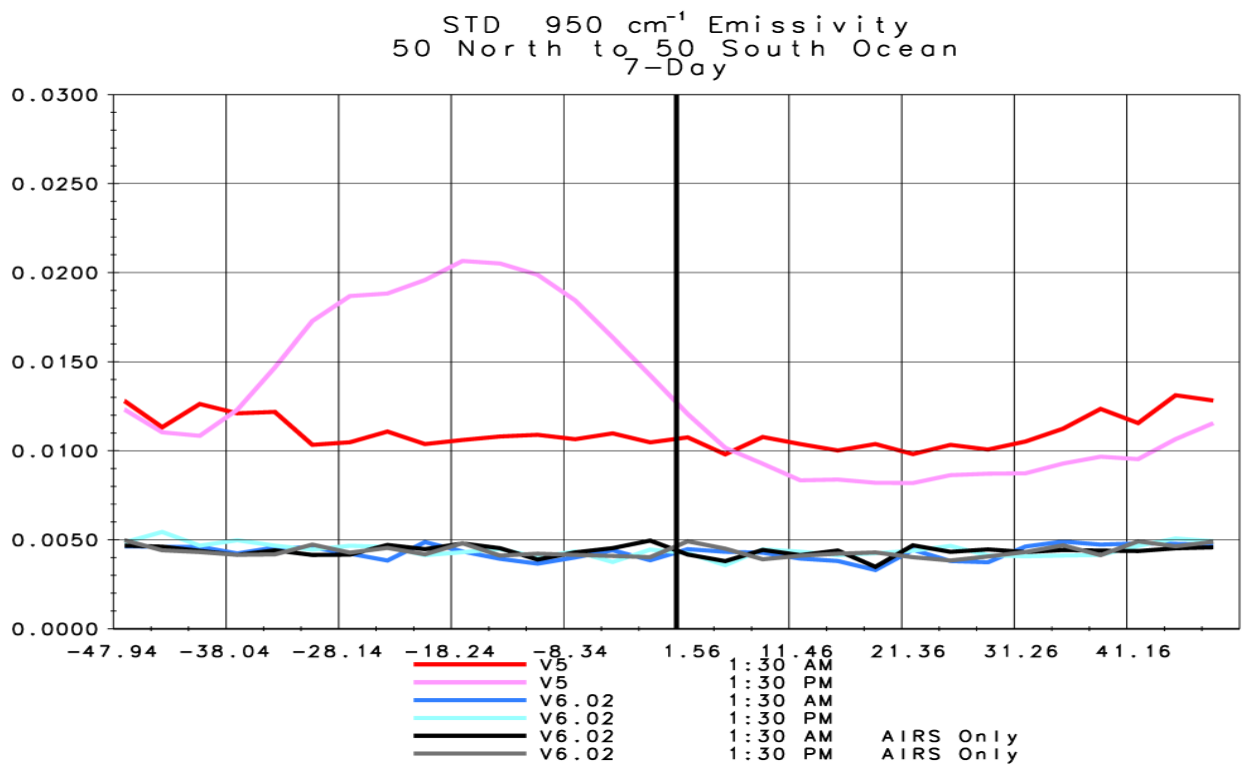


Figure 3c

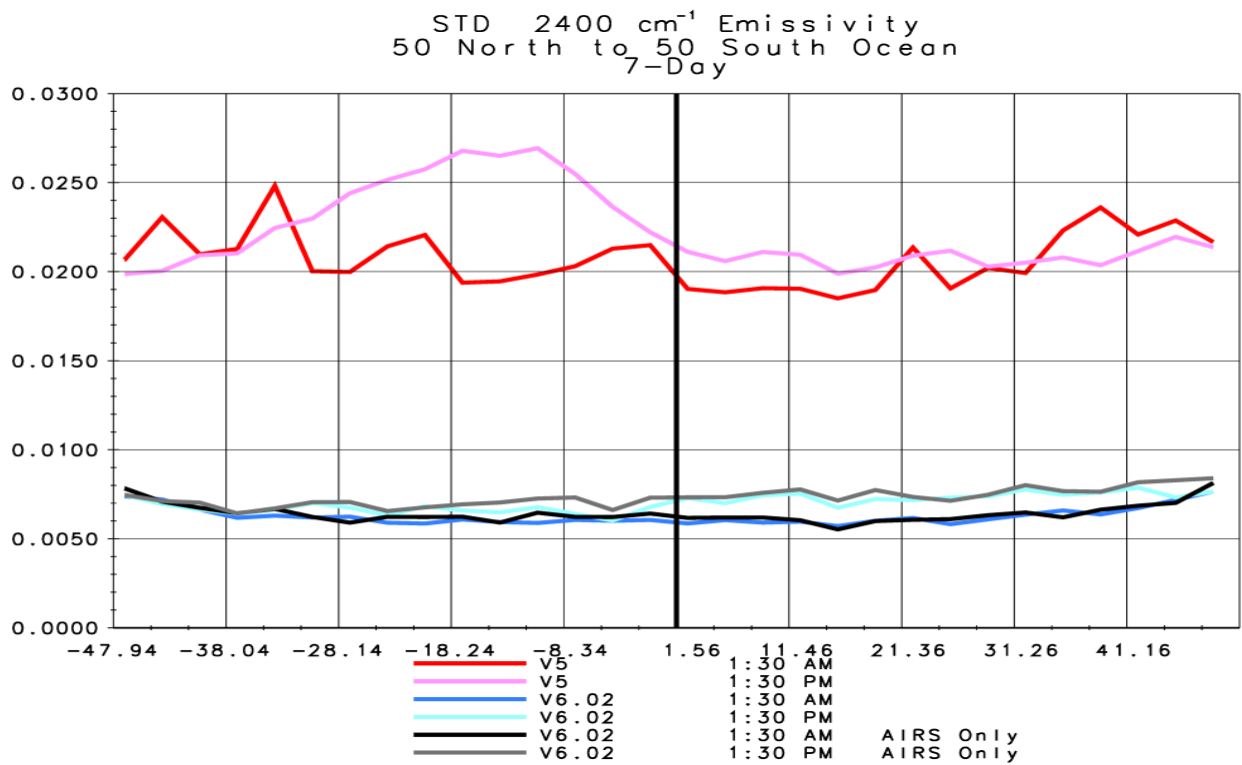


Figure 3d

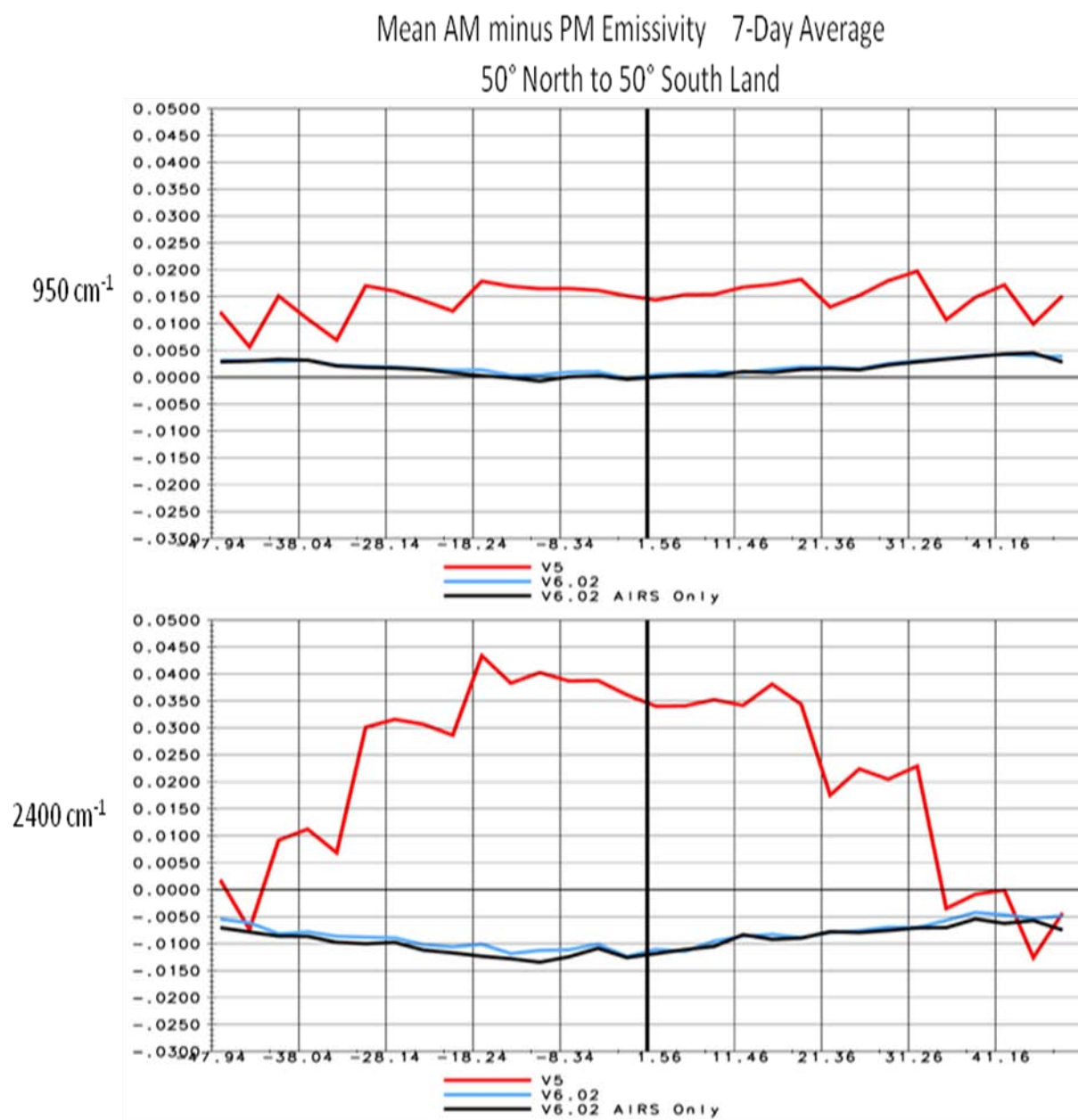


Figure 4

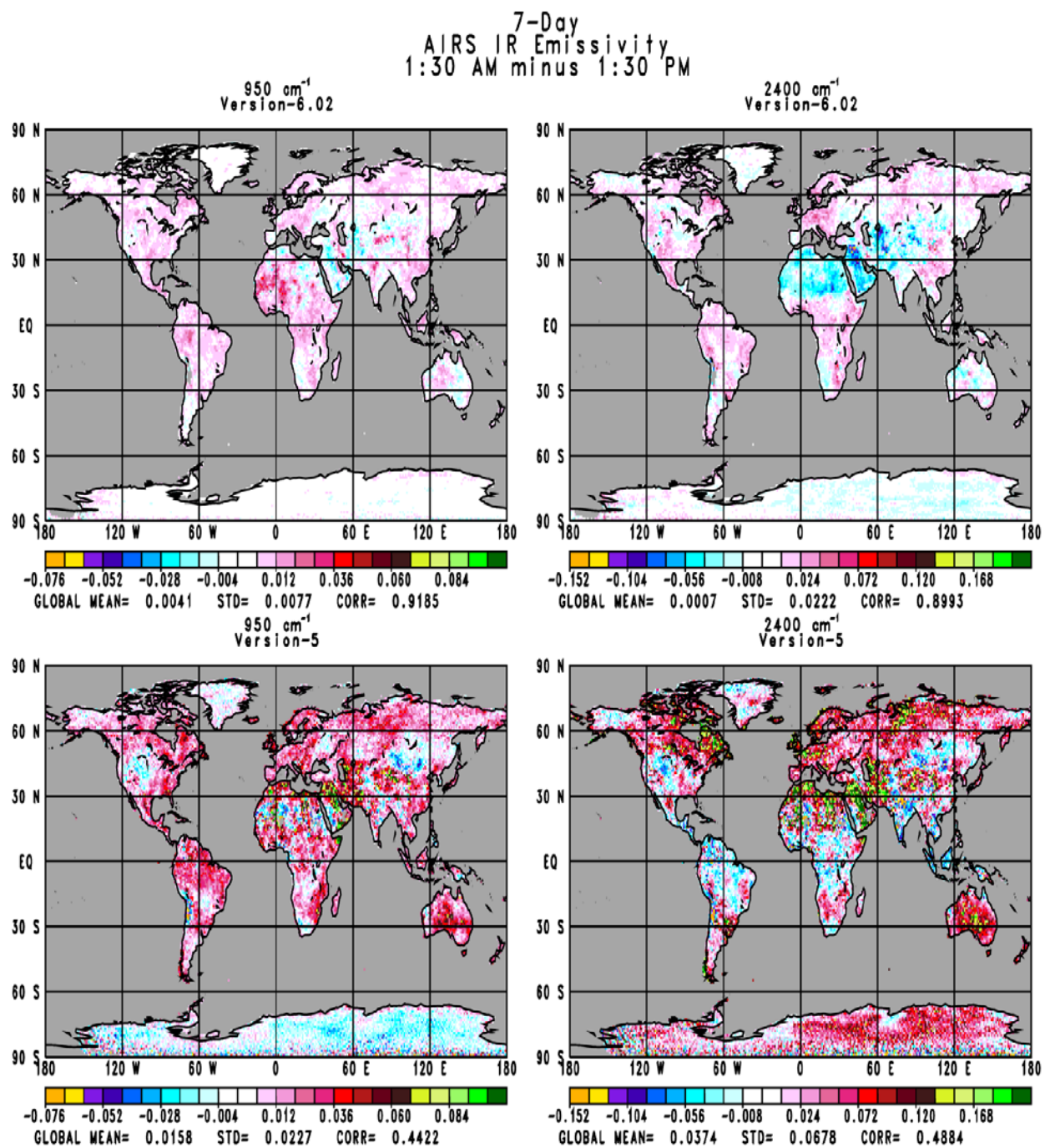


Figure 5

2.2 Temperature and Water Vapor Profiles

The fundamentals of the methodology used in Version-6 to retrieve temperature profile $T(p)$, and water vapor profile $q(p)$, from AIRS cloud cleared radiances $\tilde{R}_c$ are basically the same as those used in Version-5. Quality Controlled (QC'd) Version-6 retrievals of $T(p)$ and $q(p)$ are significantly better than those of Version-5 for three reasons: 1) Version-6 uses Neural Net generated first guesses for $T(p)$ and $q(p)$ in place of the regression generated first guess used in Version-5. The Neural Net first guesses are more accurate than the regression guesses, especially under more cloudy conditions. This allows for the generation of accurate QC'd Version-6 retrievals under cloudier cases than was achievable in Version-5. 2) Version-6 has improved QC procedures for $T(p)$ and $q(p)$ than were used in Version-5. 3) Improved Version-6 surface parameters also allow for improved Version-6 $T(p)$ and $q(p)$ in the boundary layer, especially over land.

The next section shows the improvement of Version-6 $T(p)$ and $q(p)$ retrievals compared to Version-5 in terms of accuracy and yield and also improvement of Version-6 retrievals compared to Version-5 in terms of yield and bias trends.

2.2.1 $T(p)$ and $q(p)$ Retrieval Accuracy as a Function of Yield

Figure 6 shows statistics of the differences of QC'd Version-5 and Version-6 retrievals from collocated ECMWF truth for a globe ensemble of cases taken over the 7 focus days. Panel (a) shows the percentage of QC'd cases accepted as a function of height, panel (b) shows RMS differences of 1 km layer mean temperatures from collocated ECMWF truth, and panel (c) shows biases of QC'd 1 km layer mean differences from ECMWF. Statistics are shown for six sets of results. Results for Version-5 retrievals are shown in red, results for Version-6 retrievals (called V6.02) are shown in blue, and results for Version-6 AIRS Only retrievals (called V6.02 AO) are shown in black. Version-5 did not have QC'd AIRS Only retrievals. Two sets of curves are shown for each experiment, each using different QC thresholds. Version-5 had only one set of QC thresholds, called standard thresholds. These Version-5 thresholds were chosen so as to provide a middle ground between the highest accuracy, which would be optimal for Data Assimilation purposes, and the highest yield (best spatial coverage), which would be optimal for Climate purposes. Experience using Version-5 products showed that Standard QC thresholds

were optimal for neither purpose. For example, Data Assimilation experiments using Version-5 retrievals that passed a tighter set of QC thresholds than found in the official Version-5 system, resulted in significantly improved forecasts compared to those passing the looser Standard QC thresholds. The solid red lines in Figure 6, and subsequent figures, shows statistics of Version-5 retrievals passing the tighter QC threshold, which we refer to as Tight QC threshold, and the dashed red lines show equivalent statistics for the ensemble of Standard Version-5 retrievals. Version-6 uses two different sets of thresholds, a very tight set of thresholds newly optimized for Data Assimilation purposes (QC=0), and a substantial looser set of thresholds optimal for Climate purposes (QC=1). As with Version-5, the solid lines show V6 and V6 AO results using the Data Assimilation (DA) QC thresholds, and the dashed lines show results using the Climate thresholds. Level-3 gridded products utilize all cases passing Climate QC.

In Version-5, all retrievals are either accepted or rejected above 70 mb based on use of different types of tests, even before the QC procedures are applied. One of the tests that eliminates consideration of the entire temperature profile, and flags the entire profile with QC=2 (do not use), is that the retrieved cloud fraction is over 90%. Roughly 83% of Version-5 retrievals pass the initial screening procedure, but none of them are in near overcast conditions. Version-5 retrievals with Tight QC have considerably lower yield than those with Standard QC below 200 mb, with correspondingly smaller RMS errors, on the order of 1K beneath 300 mb. There is no appreciable difference in Version-5 bias errors compared to ECMWF found using either set of QC thresholds.

Version-6 does not apply any test which eliminates the entire temperature profile, other than the requirement that the retrieval runs to completion. Version-6 retrievals using DA thresholds have roughly 1K RMS errors throughout the atmosphere, with a yield which is much higher than Version-5 Tight down to about 500 mb. The yield of Version-6 retrievals with DA QC is lower than that of Version-5 Tight beneath 500 mb, but with a considerable improvement in mid-lower tropospheric temperature RMS errors, with values less than 1K, which is believed to be optimal for Data Assimilation purposes. The yield of Version-6 retrievals with Climate QC is extremely high throughout the atmosphere, with a value of about 83% at the surface. Achievement of this very high yield is extremely valuable in the generation of more

representative Level-3 products which are used for Climate data sets. RMS errors of Version-6 retrievals with Climate QC are better than, or comparable to those of, Version-5 Standard down to about 700 mb, but with a much higher yield. Beneath 700 mb, Version-6 Climate QC RMS errors are somewhat larger than those of Version-5 Standard, but the Version-6 results are essentially unbiased, which is the more important statistic with regard to the generation of the Level-3 products used for Climate research. QC'd results for Version-6 AO are roughly comparable to those of Version-6 but with a somewhat lower yield near the surface.

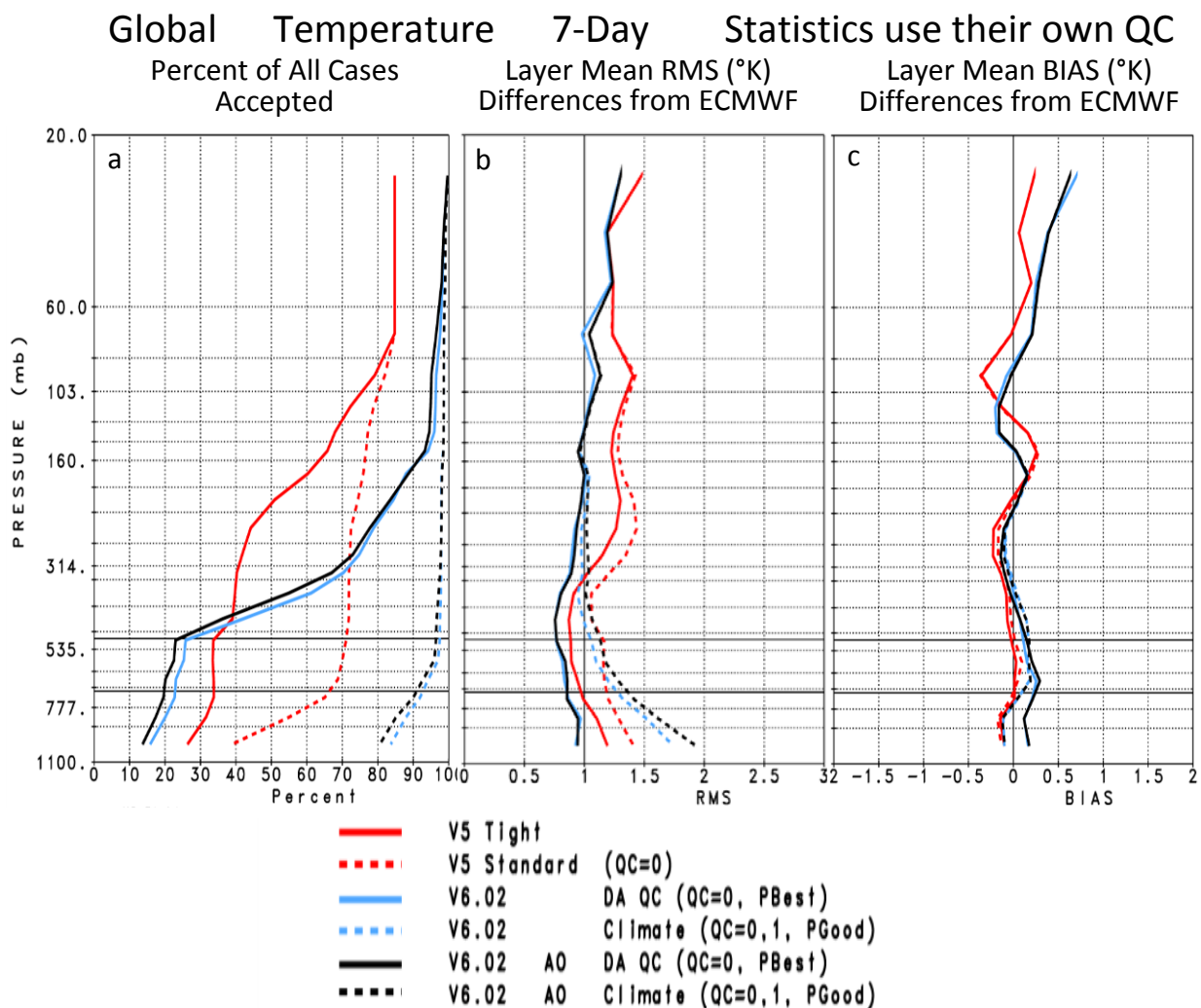


Figure 7 shows analogous results comparing QC'd 1 km layer precipitable water to that of collocated values of ECMWF. Figure 7 contains results for only Version-5 retrievals and Version-6 retrievals. Results are shown only up to 200 mb, above which water vapor retrievals

are considered to be of minimal validity, and are not included in the Standard Product data set. The relative results regarding Version-5 and Version-6 are analogous to those found for $T(p)$. Version-6 $q(p)$ retrievals with DA QC are considerably improved over those of Version-5 in the lower troposphere. This improvement is at least partially a result of the improved values of T_s , ϵ_v in Version-6 compared to Version-5. As with $T(p)$, Version-6 $q(p)$ retrievals with Climate QC are unbiased, have high accuracy, and contain almost complete spatial coverage.

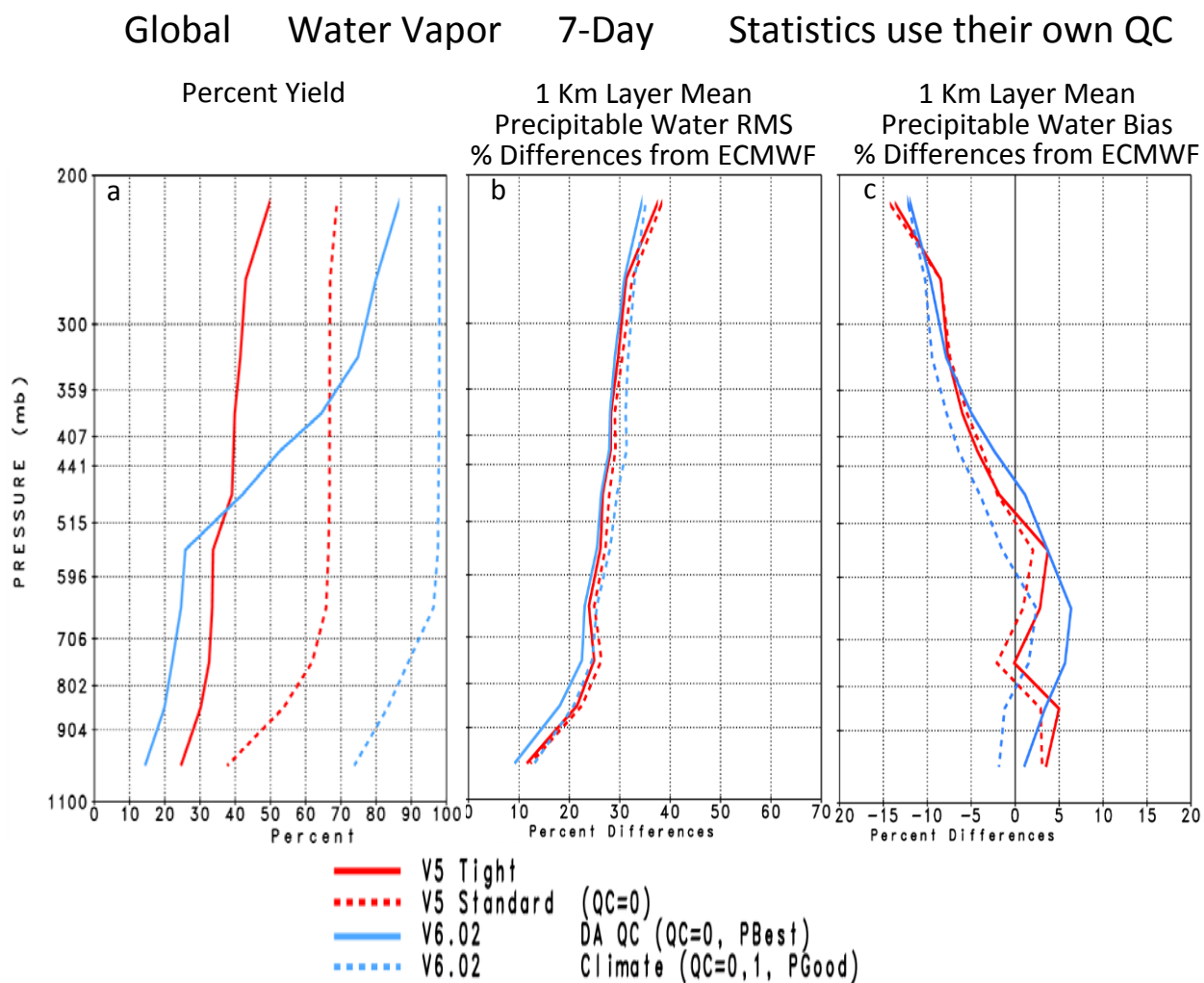


Figure 7

Figure 8 shows analogous results comparing QC'd Version-6 $q(p)$ retrievals with those of Version-6 AO. As in the case of $T(p)$, Version-6 AO water vapor retrievals are somewhat poorer than those of Version-6, but are still of very high quality. Part of this degradation results from

loss of the information in Version-6 AO contained in the two channels of AMSU-A2, which are very sensitive to boundary layer water vapor over ocean.

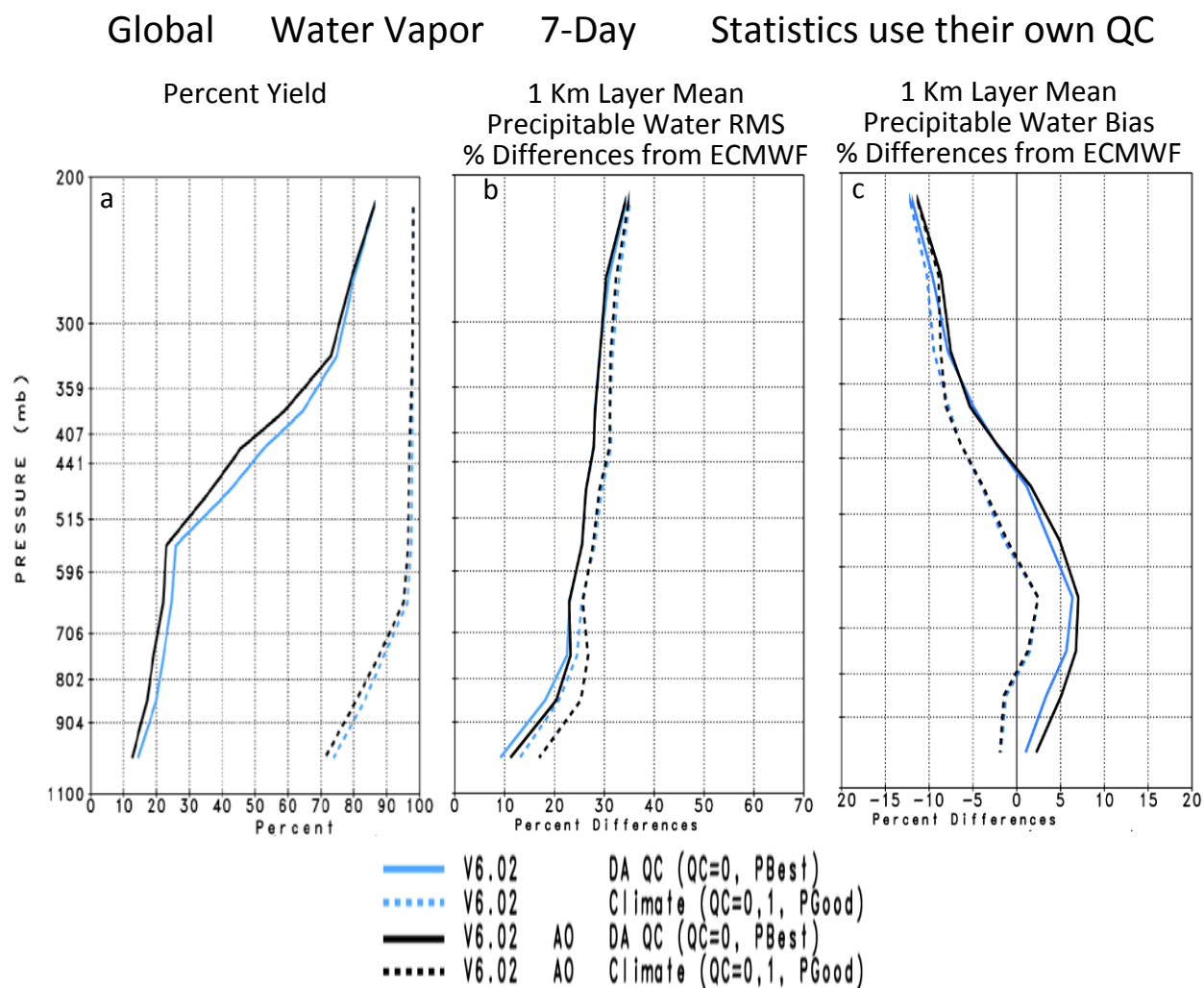


Figure 8

Figures 6-8 provide very important information about the accuracy of Quality Controlled retrievals obtained by different retrieval systems each using their own QC procedures. Indeed, the ability of a different retrieval system to perform QC is a critical part of that retrieval system, especially in the generation of Level-3 products. Figures 6-8 do not tell the whole story about the relative accuracy of the retrievals obtained in Version-5 and Version-6 however, because results are shown for different ensembles of cases. Figure 9 compares RMS T(p) errors of Version-6 and Version-5 retrievals when evaluated on common ensembles of cases. Results for two such ensembles are shown: an ensemble of relatively easier (less cloudy) cases given by

those cases accepted in Version-5 using Tight QC (shown in solid lines); and an ensemble including much more difficult (more cloudy) cases given by those cases accepted in Version-6 using Climate QC, shown in dashed lines. As previously, Version-6 RMS errors are shown in blue and Version-5 RMS errors are shown in red. Version-6 retrievals for the easier (solid line) cases are more accurate than those of Version-5 at all levels, but the degree of improvement below 500 mb is relatively small for these cases. The accuracy of Version-5 retrievals degrades much more rapidly than those of Version-6 for the harder cases (dashed lines). In fact, it is for this reason that the Version-5 retrieval system did not use relaxed QC thresholds that would have provided much for higher yields to be used in the generation of Level-3 products.

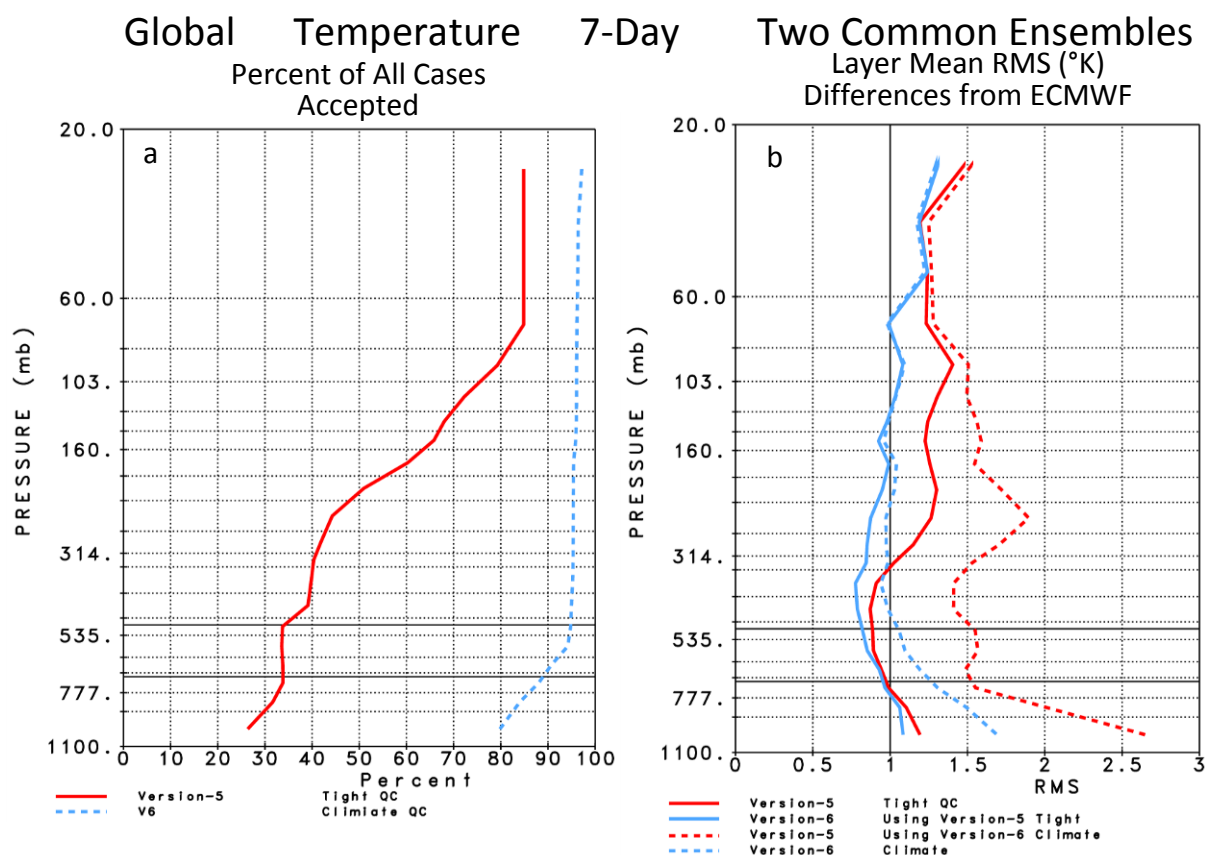


Figure 9

All results shown so far have been for a global ensemble of cases. Table 1 contains a breakdown of two temperature profile statistics, the Tropospheric Temperature Metric (TTM) and the Boundary Layer Metric (BLM), evaluated over different spatial regions: global; land 50°N to

50°S; ocean 50°N to 50°S; poleward of 50°N; and poleward of 50°S. In this table, TTM represents the mean RMS T(p) error over all 1km layers from the surface to 100 mb, and BLM represents the mean RMS T(p) error over the 6 lowest 0.25 km layers from the surface. TTM and BLM results for Version-5 and Version-6 retrievals evaluated over the Version-5 Tight ensemble are shown in Table 1a, and evaluated over the Version-6 Climate QC ensemble are given in Table 1b.

Table 1. 7-Day Mean Statistics Tropospheric Temperature Metric (TTM) and Boundary LayerMetric (BLM)

1a. Cases in Common Using the Version-5 Tight Ensemble

	<u>Global</u>		<u>Land ± 50</u>		<u>Ocean 50</u>		<u>Poleward of 50°N</u>		<u>Poleward of 50°S</u>	
	TTM	BLM	TTM	BLM	TTM	BLM	TTM	BLM	TTM	BLM
Version-5	1.10	1.29	1.19	1.71	1.04	1.13	1.14	1.50	1.31	1.76
Version-6.02	0.92	1.16	0.94	1.49	0.86	0.98	0.96	1.47	1.20	1.69

1b. Cases in Common Using the Version-6.02 Climate Ensemble

	<u>Global</u>		<u>Land ± 50</u>		<u>Ocean 50</u>		<u>Poleward of 50°N</u>		<u>Poleward of 50°S</u>	
	TTM	BLM	TTM	BLM	TTM	BLM	TTM	BLM	TTM	BLM
Version-5	1.67	2.57	1.82	2.78	1.65	2.48	1.53	2.39	1.72	2.72
Version-6.02	1.11	1.67	1.06	1.75	1.03	1.34	1.12	1.93	1.32	2.02

Version-6 T(p) retrievals are superior to those of Version-5 with regard to both metrics in all spatial regions. It is important to note the improvement of Version-6 Boundary Layer Temperatures compared to Version-5 especially over land, even for the easier ensemble of cases. This improvement over land is at least in part a result of the improved values of land surface temperature and surface spectral emissivity over land. Improvement of boundary layer temperatures was one of stated goals of Version-6, which indeed has been accomplished. The improvement of boundary layer temperatures is even more pronounced when evaluated over the ensemble of much more difficult cases.

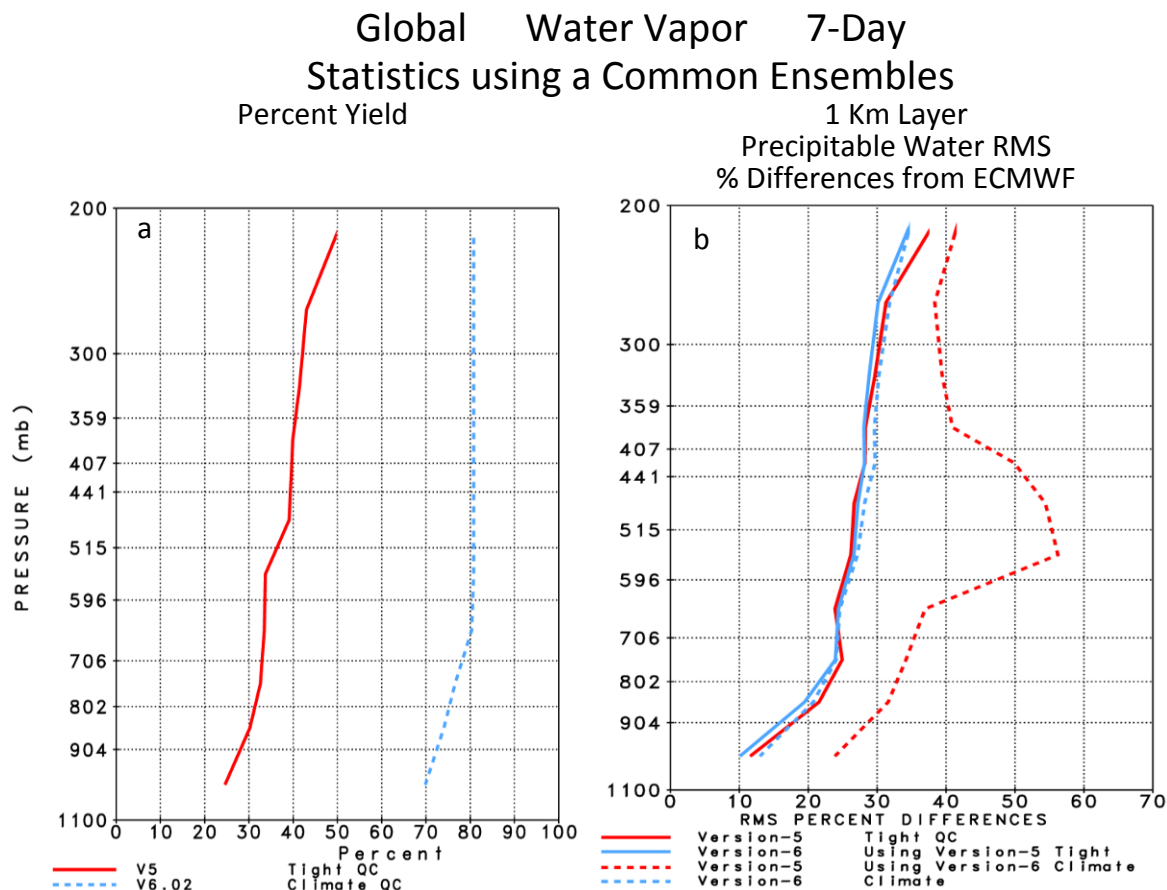


Figure 10

Figure 10 is analogous to Figure 9, but shows yields and RMS differences from ECMWF of 1km layer precipitable water for Version-5 and Version-6 retrievals evaluated over the same common ensembles of cases used for T(p) and shown in Figure 9. As in the case of T(p), layer integrated lower tropospheric water vapor is more accurate in Version-6 than in Version-5 for the easier sets of cases, and layer integrated tropospheric water vapor in all layers is much more accurate in Version-6 for the more comprehensive ensemble of cases passing Version-6 Climate QC thresholds.

Figure 11 shows the spatial distribution of a pseudo-Level-3 seven day field of accepted cases of total precipitable water, W_{TOT} , flagged to be of climate quality (QC=0,1). We refer to this spatial distribution as a pseudo-Level-3 product because the seven days are not contiguous in time. In Version-6, W_{TOT} is flagged to be of climate quality if the water vapor profile passes

the climate QC test down to the surface. Version-5 uses a different procedure to determine if W_{TOT} is of climate quality. The values shown for Version-6 and Version-5 represent the ensemble mean difference, for all accepted cases within that grid box, of the retrieved value of W_{TOT} from the collocated ECMWF value of W_{TOT} . Grid points in which no accepted values of W_{TOT} were found for any of the seven days, either daytime or nighttime, are shown in gray. Statistically, Version-6 seven day mean values of W_{TOT} are considerably more accurate than those of Version-5, both with regard to global mean bias and as well as to the standard deviation of the errors. Even more important from the climate perspective, spatial coverage of the seven day mean Version-6 product, with 99.89% of the grid boxes filled, is much more complete than that of Version-5, with 96.12% of grid boxes filled. Moreover, the 3.88% of grid boxes for which no accepted soundings were generated in Version-5 tend to come in spatially coherent groups in oceanic areas where low clouds tend to exist.

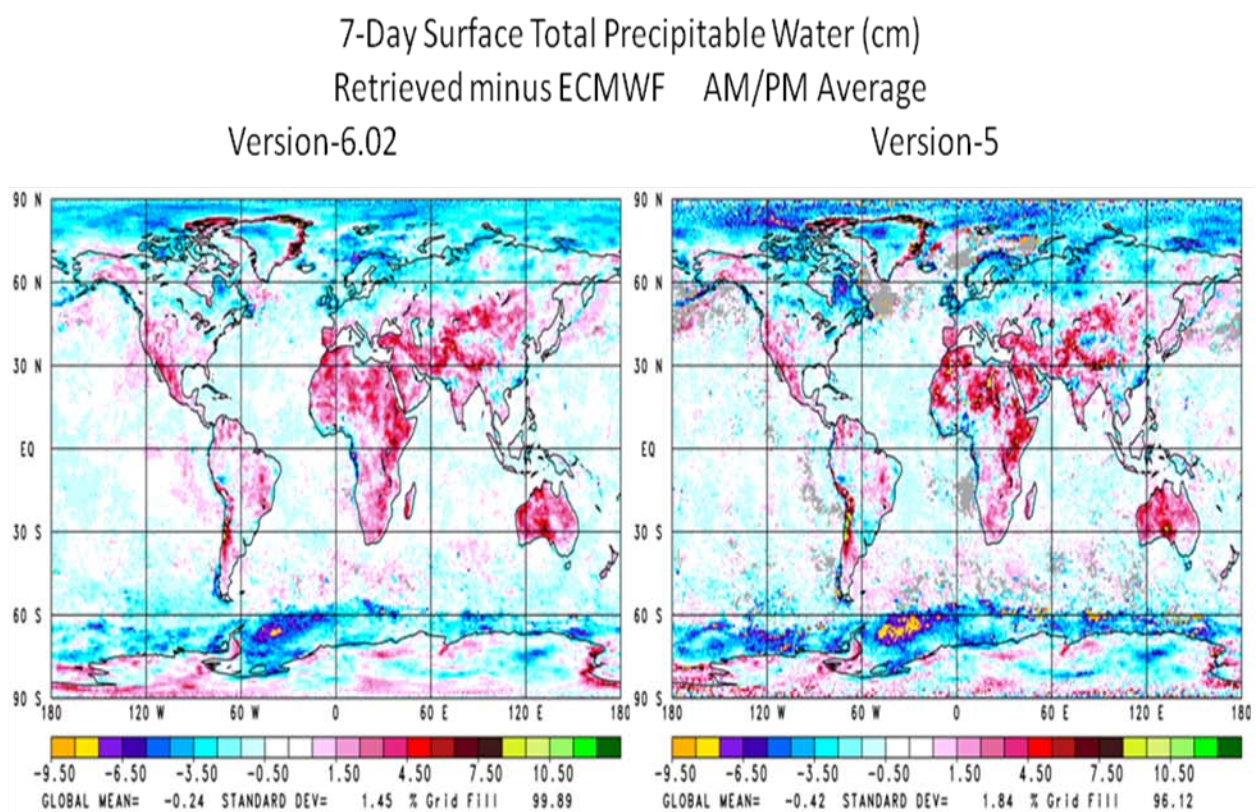


Figure 11

2.2.2 Yield and Spurious Bias Trends of $T(p)$ and $q(p)$

Version-5 retrievals have been found to have two very undesirable characteristics with regard to trends, both of which were considered to be critical to correct in Version-6. The first was that the % yield of accepted retrievals was found to be decreasing over time (negative yield trend). The second was that the mean difference of QC'd retrieved temperatures from collocated truth values were found to be changing over time as well, especially beneath 300 mb (spurious bias trend). In this section, we examine yield and spurious bias trends of Version-5 and Version-6 products. We computed volumes of these trends by taking the slope of the linear least squares fit passing through the values of the appropriate parameter for each of the seven days as a function of time.

Figure 12 shows these trends as a function of pressure for Version-5, Version-6, and Version-6 AO. All results are for cases with Climate QC (Standard QC for Version-5) because trends are most significant with regard to the generation of Level-3 products used for climate research. It is apparent that Version-6 has eliminated the substantial negative tropospheric temperature profile yield trends, on the order of 1% per year, which were found in Version-5. In addition, the Version-6 negative $T(p)$ bias trends beneath 500 mb are much smaller than those of Version-5, which were on the order of -0.05K/yr. The negative $q(p)$ bias trends found in Version-5 are also substantially reduced in Version-6. It is interesting to note that the spurious $q(p)$ bias trends found in Version-5, as a function of pressure, tend to follow those of $T(p)$ in sign. This is consistent with physics in that a spuriously cold temperature solution (trend) result lowers computed radiance for water vapor channels. If the retrieved temperature is too low, the solution for $q(p)$ will result in a lowered retrieved water amount so as to raise the radiances computed using the incorrect value of $T(p)$ in order to match the observed radiance, resulting in a spurious negative low value (or negative trend) of $q(p)$. This lowered $q(p)$ (trend) subsequently gives too high a computed radiance in the window channels used to compute cloud parameters, which subsequently results in increased values of retrieved cloud fraction, and finally a spurious increasing cloud trend.

Global Trends 7-Day

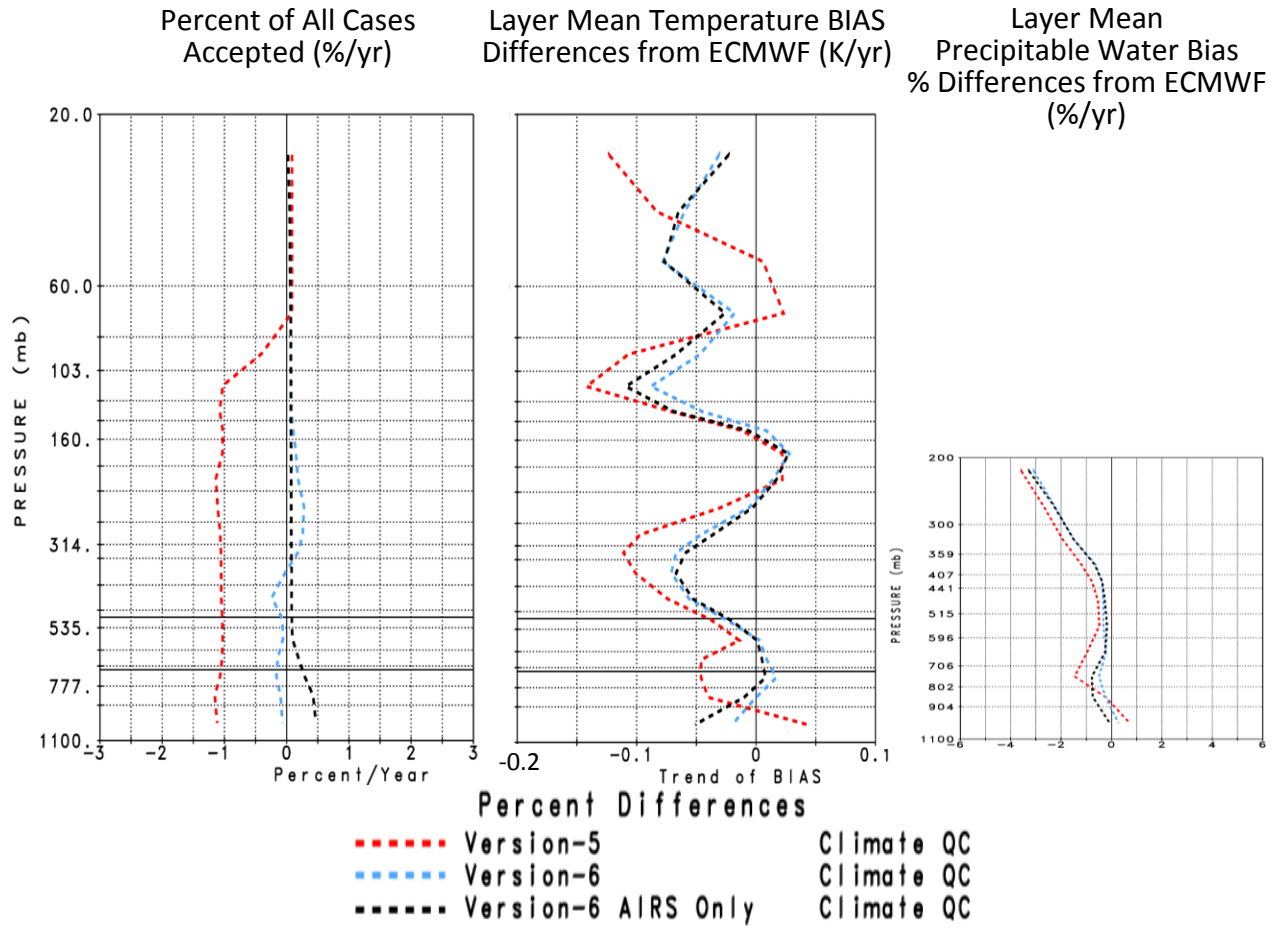


Figure 12

2.3 Quality controlled values of clear column radiances $\hat{R}_i$

The clear column radiance for channel i , $\hat{R}_i$, is a derived quantity which represents the radiance channel i would have observed if the entire 3x3 array of AIRS FOV's used to generate a retrieval were completely clear. Retrievals are generated for this 3x3 array treating $\hat{R}_i$ as the observed radiance. $\hat{R}_i$, like all other derived quantities, have case-by-case, channel-by-channel error estimates $\hat{R}_i$. The accuracy of $\hat{R}_i$ can be attested to by the ability to generate accurate QC'd retrievals under most cloud conditions as discussed in Section 2.2.1. In this section, we evaluate the accuracy and yield of Version-6 Quality Controlled values of the product $\hat{R}_i$.

The top panel of Figure 13a shows the percent of all cases, as a function of frequency, of cases passing the Climate QC thresholds (light colors) and Data Assimilation QC thresholds (dark colors) respectively. Results are shown in blue for Version-6 and in black for Version-6 AO. The QC tests use the error estimate $\delta \bar{R}_i$, converted into brightness temperature units $\delta \bar{\Theta}_i$, where $\delta \bar{\Theta}_i$ is the uncertainty in channel i brightness temperature given a radiance uncertainty $\delta \bar{R}_i$ when evaluated at the clear column brightness temperature $\bar{\Theta}_i$. The QC flag is set equal to 0 if $\delta \bar{\Theta}_i \leq 1K$, set equal to 1 if $1K < \delta \bar{\Theta}_i \leq 2.5K$, and set equal to 2 if $\delta \bar{\Theta}_i > 2.5K$. The second panel shows the mean value of $\bar{\Theta}_i$ taken for all cases with QC ≤ 1 . The third and fourth panels show statistics for the standard deviation of QC'd values of $(\bar{\Theta}_i - \Theta_i^{truth})$ and mean values of $(\bar{\Theta}_i - \Theta_i^{truth})$, respectively. The third panel also shows in yellow the mean values of the channel noise $NE\Delta T_i$ evaluated using the channel radiance noise $NE\Delta T_i$ evaluated at $\bar{\Theta}_i$.

The value Θ_i^{truth} , against which $\bar{\Theta}_i$ is evaluated, has some uncertainties associated with it. Θ_i^{truth} is computed on a case-by-case basis using the collocated ECMWF state. Errors in both the state and the Radiative Transfer Algorithm (RTA) used to compute Θ_i^{truth} will each contribute to errors in Θ_i^{truth} . Over land, the ECMWF surface parameters T_s and ϵ_v both contain considerable uncertainty. In addition, ECMWF values of water vapor and O_3 profiles carry considerable uncertainty, as well as ECMWF values of $T(p)$ in the upper stratosphere. For these reasons, the validation of $\bar{\Theta}_i$ is most reliable in those channels which are relatively insensitive to the surface, temperatures above 10 mb, water vapor, or ozone. In addition, observed radiances at frequencies higher than 2180 cm^{-1} are sensitive in varying degrees to the effects of solar radiation reflected by the surface. ECMWF does not contain a value of the surface bi-directional reflectance, so effects of reflected solar radiation are not included in the computations of Θ_i^{truth} . For these reasons, the results shown in Figure 13a are for night-time ocean cases only. Figure 13b is a blow-up of Figure 13a, shown only for spectral interval 650 cm^{-1} to 760 cm^{-1} .

Figure 13a shows that yields of accepted values of $\bar{R}_i$ using climate QC thresholds are 50% or higher throughout the spectrum. Yields are higher for those channels sensing higher in the atmosphere, in which observed radiances are less sensitive to cloud cover. Window

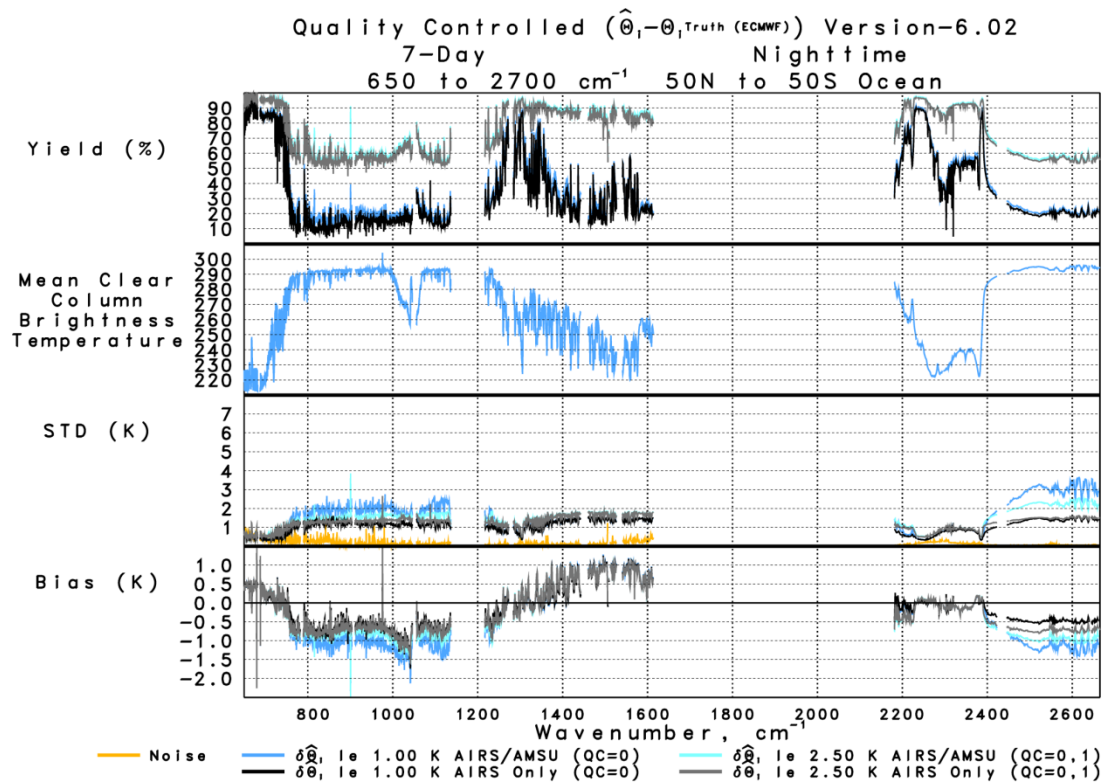


Figure 13a

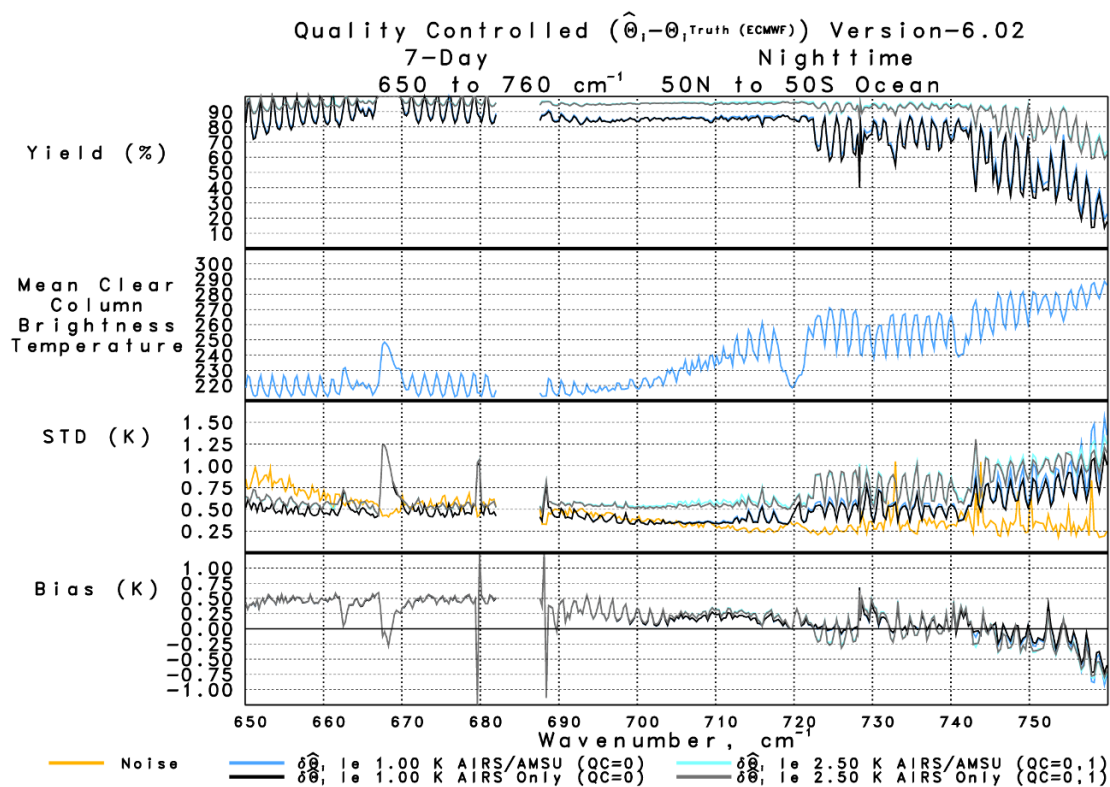


Figure 13b

channel yields over ocean are on the order of 50%, which is roughly consistent with the yield of accepted ocean surface temperatures using climate QC thresholds shown in Figure 1. The percent yields of $\hat{R}_i$ in Version-6AO are very similar to those in Version-6 at all frequencies with either set of QC thresholds. The standard deviation of $|\hat{\Theta}_i - \Theta_i^{truth}|$ are considerably larger in window regions than in the longwave and shortwave temperature sounding regions, $650\text{ cm}^{-1} - 740\text{ cm}^{-1}$ and $2180\text{ cm}^{-1} - 2390\text{ cm}^{-1}$ respectively. Part of this result is an artifact resulting from the effect of the uncertainty in ocean surface skin temperature and ocean spectral emissivity on the values of Θ_i^{truth} in channels sensitive to the surface. Part of this increase in apparent “error” in $\hat{\Theta}_i$ may well be a real result because window channel radiances are more sensitive to very low clouds over the ocean which are more difficult to handle precisely in the generation of $\hat{R}_i$. Biases $\hat{\Theta}_i - \Theta_i^{truth}$ for all 4 ensembles of cases are similar to each other in all spectral regions except for the window regions. These biases outside the window regions are more likely due to biases in Θ_i^{truth} rather than $\hat{\Theta}_i$ and are a result of systematic errors in the RTA as well as in the ECMWF truth vector. The negative values of $\hat{\Theta}_i - \Theta_i^{truth}$ in window regions, on the order of -0.5K—1.0K, are most likely real and the result of insufficient cloud clearing especially when very low clouds are present. This is consistent with the fact that this negative bias is larger when the less restrictive climate QC threshold is used.

The most important potential application of Quality Controlled values of $\hat{R}_i$ is with regard to data assimilation. ECMWF and NCEP assimilate AIRS data as part of their operational Data Assimilation procedure. In particular, ECMWF and NCEP assimilate AIRS radiances in the spectral interval 650 cm^{-1} to 750 cm^{-1} on a case-by-case, channel-by-channel basis, for those channels whose observed radiances are thought to be unaffected by clouds. In principle, operational centers could assimilate QC’d cloud cleared radiances in an analogous way given appropriate QC procedures. The spatial coverage of QC’d cloud cleared radiances is potentially much greater than that of radiances unaffected by clouds. Figure 13b shows that accepted values of $\hat{R}_i$ with QC=0 over ocean for the most part have yields of 75% or better at frequencies less than 740 cm^{-1} . In addition, the standard deviation of the errors in $\hat{R}_i$ with QC=0 are on the order of the channel noise at these frequencies. For those channels in which the errors in $\hat{R}_i$

are greater than the channel noise, their individual errors are characterized very well by $\delta\hat{R}_i$ and this can be taken into account by the data assimilation procedure. The errors in $\Delta\hat{\theta}_i$ are actually lower than the channel noise for stratospheric sounding channels because $\hat{R}_i$ is given by the average of the 9 observed values of $\hat{R}_i$ in the 9 different FOV's used to generate a retrieval for cases thought to be unaffected by clouds.

3. Validation of Monthly Mean Level-3 Products

The previous discussion has dealt with validation of Level-2 products for an ensemble of 7 four days, in different seasons and years, using ECMWF at “truth”. This section compares Version-5 and Version-6.03 monthly mean products for the 12 months in which Version-6.03 was run: January, April, July, and October in each of 2003, 2007, and 2011. The emphasis of the validation is with regard to “trends”, or more importantly the difference in trends of Version-6 and Version-5, where the “trend” is defined as the slope of the linear least squares first passing through data of all 12 months with a linear time scale. Particular attention will be paid to trends of OLR, clear sky OLR, temperature profile, mid-tropospheric water vapor mixing ratio, total precipitable water, and fractional cloud cover.

Figure 14 is not related to trends but rather shows differences of AIRS global mean monthly mean OLR (green) and Clear Sky OLR (red) values from those generated by the CERES Science Team using CERES observations. Version-5 differences from CERES are shown for each month of the overlap OLR time series, September 2002 through October 2011, and Version-6.03 differences are shown for each of the twelve months for which Version-6.03 was run. The solid lines (green and red) are horizontal lines, passing through the mean differences between AIRS Version-5 and CERES values of OLR and Clear Sky OLR respectively, and the dashed lines (green and red) are horizontal lines passing through the mean differences of AIRS Version-6.03 OLR and Clear Sky OLR from the CERES values of these parameters. Version-5 AIRS OLR and Clear Sky OLR values are significantly biased with regard to those of CERES, with a bias that is essentially constant in time but with a small seasonal cycle. Previous work has shown that the anomaly time series of AIRS and CERES OLR, as well as of AIRS and CERES Clear Sky OLR, are in

very close agreement with each other. Nevertheless, the large biases of 8.59 W/m^2 and 7.96 W/m^2 for AIRS OLR and Clear Sky OLR respectively with respect to CERES are somewhat

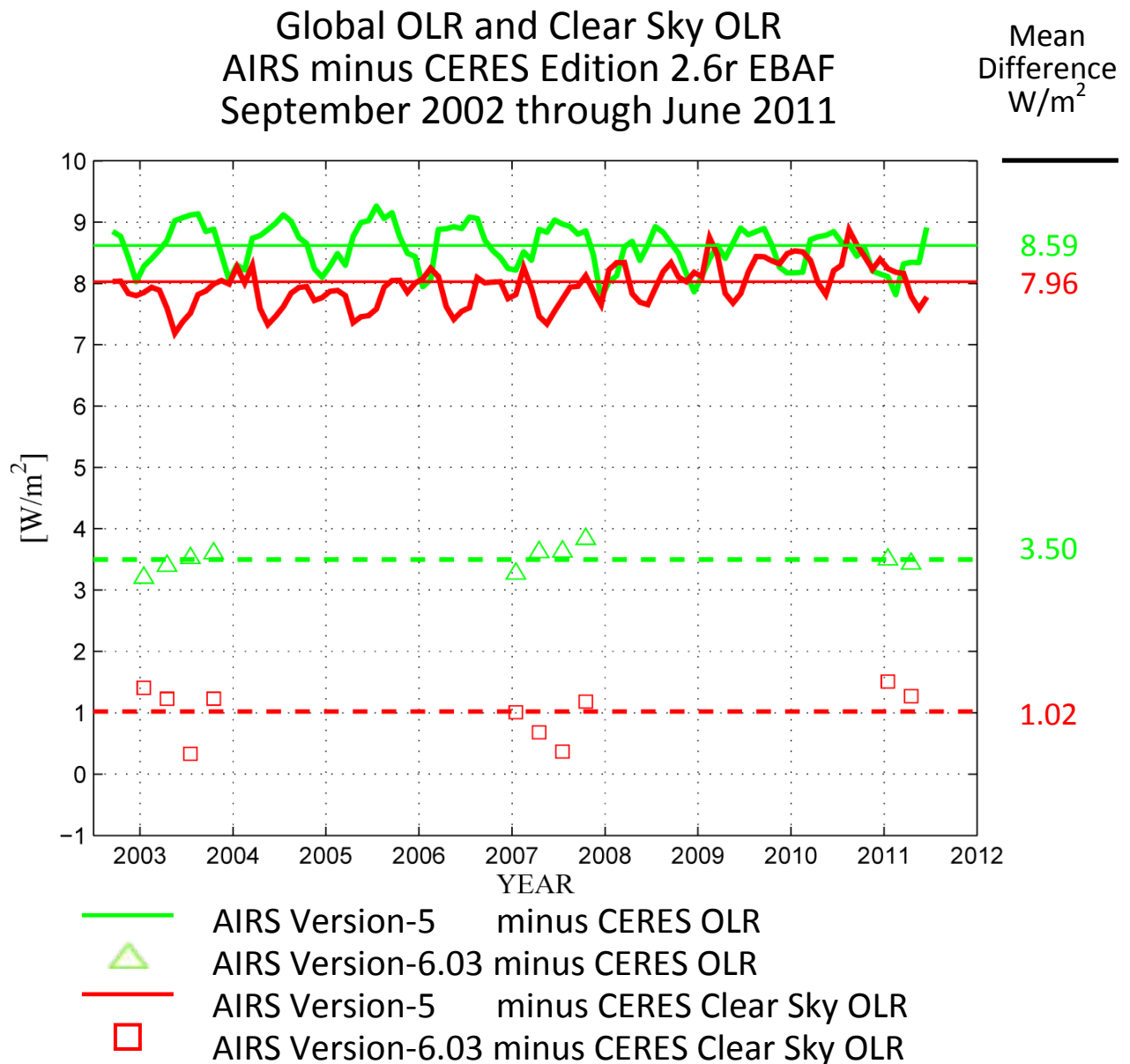


Figure 14

disconcerting. Figure 14 shows that Version-6.03 OLR and Clear Sky OLR biases compared to CERES will be reduced to much more acceptable values of 3.37 W/m^2 and 1.02 W/m^2 respectively. These differences are generally on the order of the uncertainty of the CERES measurements themselves.

Figure 15 shows values in blue of the global mean time series of six Version-5 products: OLR; Clear Sky OLR; 500 mb temperature; 500 mb water vapor mixing ratio; Total Precipitable Water; and Effective Cloud fraction. Version-5 values for each of the months between January 2003 and October 2011 are given by the dashed blue lines. Version-5 values for the twelve comparison months are also indicated by the blue stars which lie on the dashed blue curve. The solid blue lines show the slopes of the linear least squares fits passing through the 12 blue stars, which we have referred to as “trends”. Figure 15 also shows values of Version-6.03 products for the same 12 months in pink circles, with the straight lines passing those circles given in pink. Finally, the differences between Version-5 products and those of Version-6.03 are shown as black diamonds, and the straight lines passing through those differences, with an offset, are shown in black. The most important part of this figure is the slope of the black lines, which gives estimates of how actual trends of Version-6 products, when evaluated over a long time period, would differ from those of Version-5.

A number of important features are evident from the panels in this figure. While the relative slopes of the Version-5 and Version-6.03 trend lines are generally similar for OLR and Clear Sky OLR, they are considerably different from each other for the remaining geophysical parameters. Values of the slopes of all the lines are given in Table 2. For example, Version-5 500 mb temperature has a negative “trend” of -0.058 K/yr, while the “trend” in Version-6.03 is -0.006 K/yr. These “trends” can be misleading because not only does the time period used contain data from only 12 months in 3 years, but even more significantly, only a portion of the annual cycle is captured. The more significant values are the differences in the two sets of “trends” because these indicate the extent that trends of Version-6 products should differ from those of Version-5 whatever they are. For example, one would expect the trend of Version-6 500 mb temperature to be on the order of 0.052 K/yr more positive (less negative) than that of Version-5. Likewise, the Version-6 trend of 500 mb water vapor mixing ratio is expected to be less negative (or more positive) than that of Version-5 as with Total Precipitable Water. Figure 12 showed that Version-5 tropospheric temperature and water vapor had spurious negative bias trends vs. ECMWF “truth”, based on results for seven days, which were for the most part

Global Time Series January 2003 through October 2011

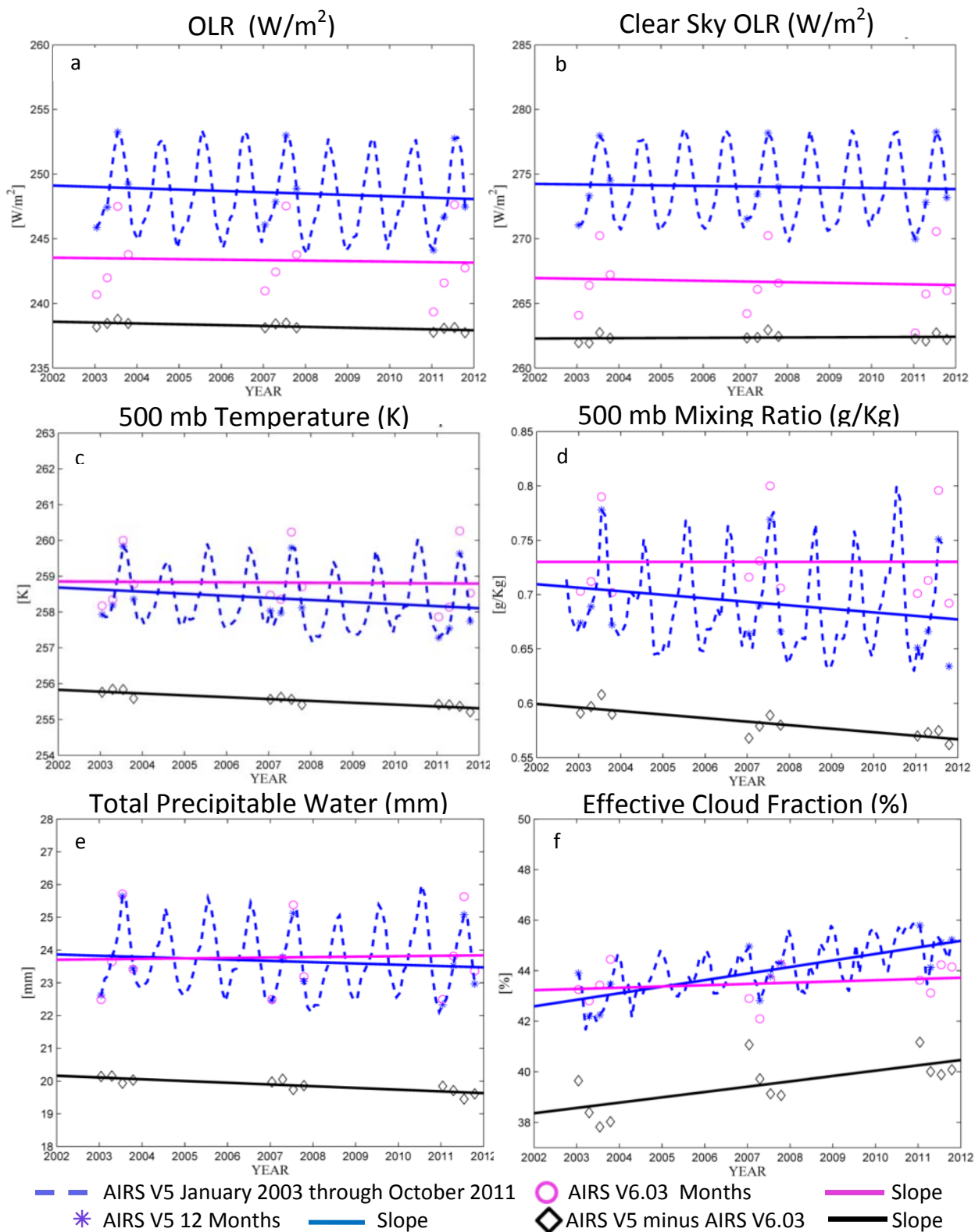


Figure 15

Table 2**12 Month Global Time Series Slopes (Trends)
January 2003 through October 2011**

	OLR W/m ² /yr	Clear Sky OLR W/m ² /yr	500 mb Temp K/yr	500 mb Mixing Ratio g/Kg/yr	Total Precipitable Water mm/yr	Cloud Fraction %/yr
* AIRS V5	-0.104	-0.040	-0.058	-0.00325	-0.0392	0.260
○ AIRS V6.03	-0.038	-0.054	-0.006	0.00001	0.0122	0.049
◇ AIRS V5 minus AIRS V6.03	-0.066	0.014	-0.052	-0.00326	-0.0514	0.211

eliminated in Version-6. Table 2 confirms this result with regard to monthly mean products over different years. Table 2 also shows that Version-5 global mean cloud cover was increasing at a most likely unreasonable rate of 0.26% per year, while Version-6 cloud fraction will increase on the order of 0.21% per year less. As discussed previously, the spurious increase in Version-5 cloud cover over time is consistent with the spurious negative trends of T(p) and q(p) found in Version-5.

Figure 16 shows trends, and trend differences related to Version-5 and Version-6 temperature profiles. The left panel in Figure 16 shows “trends” of Version-5 and Version-6 T(p) computed at different pressure levels as discussed previously, in red and blue respectively, as well the more meaningful statistic, which is the difference in the expected trends of Version-6 temperatures compared to Version-5, shown in green. The appropriate values for 500 mb temperature, given in Table 2, are plotted in Figure 16 at 500 mb. The same green curve is shown in the right panel of Figure 16. Superimposed on the green curve, showing the Version-5 trends minus Version-6 trends, is the dashed red curve which was previously shown in Figure 11. The dashed red curve shows the spurious Version-5 temperature profile bias trends as compared to ECMWF, determined from the seven focus days. To the extent that trends of Version-5 minus Version-6 temperatures match those of Version-5 minus ECMWF, then one would expect Version-6 temperature trends to match those of ECMWF very closely. The green

line, aside from being smoother in the vertical, matches the dashed red line very closely in the troposphere. This is a further confirmation that the spurious temperature trends found in Version-5 will be much smaller in Version-6 as desired.

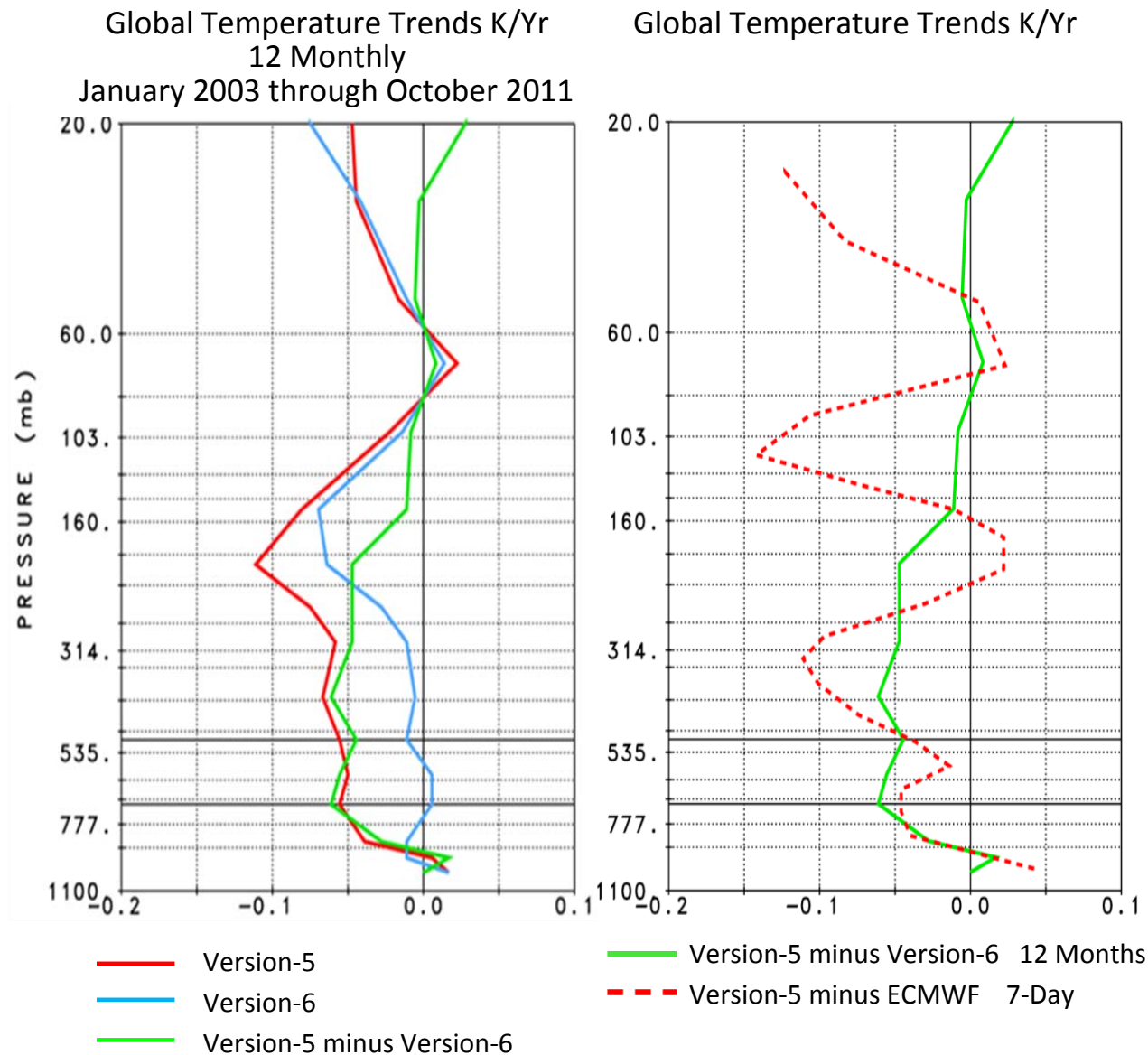


Figure 16

4. Summary

Our validation studies of Version-6 products compared to those of Version-5 indicate that, with regard to surface skin parameters, temperature and moisture profile, OLR and Clear Sky OLR, and cloud parameters, Version-6 products are superior in every way with regard to both accuracy and trends.